

Composite Impact of Global Hawk Unmanned Aircraft Dropwindsondes on Tropical Cyclone Analyses and Forecasts

HUI CHRISTOPHERSEN, ALTUG AKSOY, AND JASON DUNION

Cooperative Institute for Marine and Atmospheric Studies, University of Miami, and NOAA/Atlantic Oceanographic and Meteorological Laboratory/Hurricane Research Division, Miami, Florida

SIM ABERSON

NOAA/Atlantic Oceanographic and Meteorological Laboratory/Hurricane Research Division, Miami, Florida

(Manuscript received 20 October 2017, in final form 11 May 2018)

ABSTRACT

The impacts of Global Hawk (GH) dropwindsondes on tropical cyclone (TC) analyses and forecasts are examined over a composite sample of missions flown during the NASA Hurricane and Severe Storm Sentinel (HS3) and the NOAA Sensing Hazards with Operational Unmanned Technology (SHOUT) field campaigns. An ensemble Kalman filter is employed to assimilate the dropwindsonde observations at the vortex scale. With the assimilation of GH dropwindsondes, TCs generally exhibit fewer position and intensity errors, a better wind–pressure relationship, and improved representation of integrated kinetic energy in the analyses. The resulting track and intensity forecasts with all the cases generally show a positive impact when GH dropwindsondes are assimilated. The impact of GH dropwindsondes is further explored with cases stratified by intensity change and presence of crewed aircraft data. GH dropwindsondes demonstrate a larger impact for nonsteady-state TCs [non-SS; 24-h intensity change larger than 20 kt ($\sim 10 \text{ m s}^{-1}$)] than for steady-state (SS) TCs. The relative skill from assimilating GH dropwindsondes ranges between 25% and 35% for either the position or intensity improvement in the final analyses overall, but only $\sim 5\%$ – 10% for SS cases alone. The resulting forecasts for non-SS cases show higher skill for both track and intensity than SS cases. In addition, the GH dropwindsonde impact on TC forecasts varies in the presence of crewed aircraft data. An increased intensity improvement at long lead times is seen when crewed aircraft data are absent. This demonstrates the importance of strategically designing flight patterns to exploit the sampling strengths of the GH and crewed aircraft in order to maximize data impacts on TC prediction.

1. Introduction

The National Oceanic and Atmospheric Administration (NOAA)'s unmanned aircraft system (UAS) program's Sensing Hazards with Operational Unmanned Technology (SHOUT; 2015–16) was designed to collect observations in high-impact weather events such as tropical cyclones (TCs). One such UAS that the program has extensively utilized is the Northrup–Grumman Global Hawk (GH; Watts et al. 2012). The GH operates at an altitude range of 16 750–19 800 m (55 000–65 000 ft), has a cruising speed of 620 km h^{−1} (~ 335 kt), and can fly a range over 17 000 km (9000 n mi). It was previously used to collect TC observations by NASA during the Genesis and Rapid Intensification Processes (GRIP;

2010) experiment (Braun et al. 2013) and the Hurricane and Severe Storm Sentinel (HS3; 2012–14; Braun et al. 2016).

One of the goals of SHOUT is to assess the influence of UAS data on the prediction of high-impact weather events such as TCs, particularly those affecting the United States. The present study performs a set of data denial experiments to evaluate the impact of GH datasets. Specifically, we focus on GH dropwindsondes (dropsondes hereafter; Hock and Franklin 1999) from both HS3 and SHOUT field campaigns.

UAS platforms are capable of collecting observations in dangerous weather conditions that are otherwise not reachable by crewed aircraft (Cione et al. 2016). In particular, a high-altitude, long-endurance UAS, such as the GH, can provide data complementing traditional TC reconnaissance and/or surveillance aircraft. For example, while NOAA's WP-3D Orion (P-3) and Gulfstream IV-SP

Corresponding author: Hui Christophersen, hui.christophersen@noaa.gov

TABLE 1. Summary of the TC names, number of cases per TC, peak intensity at the available analysis times, and number of cases per TC when aircraft reconnaissance and/or surveillance are present.

TC name	No. of cases	No. of cases when crewed aircraft present	Peak intensity (kt)
Nadine (2012)	6	0	70
Gabrielle (2013)	1	1	30
Humberto (2013)	2	0	40
Cristobal (2014)	4	1	75
Edouard (2014)	10	2	95
Erika (2015)	1	1	40
Gaston (2016)	2	0	75
Karl (2016)	5	3	60
Hermine (2016)	5	5	70
Matthew (2016)	4	1	105

(G-IV) aircraft typically operate at altitudes of 10 000 ft (~ 3000 m) and 45 000 ft ($\sim 13\,000$ m), respectively, for up to ~ 8 – 10 h of flight time (Aberson and Franklin 1999; Rogers et al. 2006), the GH operates at much higher altitudes, with an endurance of up to 24 h (e.g., Christophersen et al. 2017). Indeed, recent observational studies have demonstrated that high-altitude dropsondes can reveal the evolution of a TC's kinematic and thermodynamic structure in a much more vertically complete sense, compared to traditional TC reconnaissance aircraft (Zawislak et al. 2016; Rogers et al. 2016; Doyle et al. 2017).

Although numerous studies have shown that assimilating dropsondes results in positive impacts on TC track (Aberson and Franklin 1999; Chou et al. 2011; Majumdar et al. 2013) and intensity (Aberson and Franklin 1999; Torn 2014) forecasts, studies that focus on vortex-scale assimilation of high-altitude (i.e., 55 000–65 000 ft) dropsondes have so far been limited, largely due to the only recent availability of these high-altitude dropsondes since the GRIP experiment. The two case studies of Edouard (2014) presented in Christophersen et al. (2017) illustrated the benefits of assimilating GH high-altitude dropsondes in a high-resolution vortex-scale data assimilation (DA) system. The TC was at hurricane intensity in both cases. Specifically, in the first case, when the TC inner core was exclusively sampled by GH dropsondes, consistently better intensity forecasts were obtained than without the data, while in the second case, when GH dropsondes sampled the TC's near environment more exclusively than the inner-core region, persistent track forecast improvement was achieved. Meanwhile, Christophersen et al. (2017) also pointed out some of the challenges of assimilating inner-core GH dropsondes to represent complex storm structures, such as double eyewalls. The present study extends these case studies to a composite

analysis and seeks to generalize the impact of GH dropsondes on TC track, intensity, and structure.

While there has been some improvement in TC intensity forecasts over the past few decades (DeMaria et al. 2014), intensity change prediction still remains challenging. Improved prediction of rapid intensification (RI) continues to be a high priority for operations (Rappaport et al. 2012). Therefore, we examine the impact of assimilating GH dropsondes with respect to whether a storm is in a steady state (SS) or not (i.e., intensifying or weakening). This study is presented as follows. Section 2 discusses the cases examined, the DA and forecast systems, and the experiment setup. Section 3 shows the GH dropsondes' performance in DA observation space. The impacts of GH dropsondes on both analyses and forecasts are further investigated with all cases (section 4), cases stratified by intensity change (section 5), and cases that included additional observations from crewed aircraft (section 6). Two cases that have large track improvement with GH dropsondes are briefly discussed in section 7. A summary and concluding remarks are given in the final section.

2. Data and experiments

a. Dataset description

Table 1 lists the TCs examined in this study, along with their peak intensity at the analysis time: 10-m maximum 1-min sustained wind speed in kt ($1 \text{ kt} = 0.5144 \text{ m s}^{-1}$), as reported in the National Hurricane Center (NHC) best track dataset (Jarvinen et al. 1984). This list contains 10 TCs from the 2012–16 Atlantic hurricane seasons that were observed by GH dropsondes. The selected TCs cover a wide range of sizes, structures, and intensities and include two major hurricanes: Edouard (2014) and Matthew (2016). We perform DA experiments whenever GH dropsonde measurements are available and denote each forecast initialized at 6-hourly synoptic times as a “case.” For each TC, we list the number of cases having GH dropsondes within a $10^\circ \times 10^\circ$ assimilation domain centered on the TC and within ± 3 h of the synoptic time (Table 1). The locations and intensities at the analysis times in Table 1 are shown in Fig. 1. Five of the TCs are from the HS3 field campaign, while the rest are from the SHOUT field campaign. Because of different priorities of the two field campaigns, most of the HS3 TCs are in the eastern North Atlantic and were unreachable by crewed reconnaissance aircraft from typical U.S. deployment sites, while the SHOUT cases are closer to land.

The TC cases are further stratified by intensity change following the convention of Rogers et al. (2013). The 12 h following each analysis time are examined to include the most relevant vortex- and convective-scale processes. A

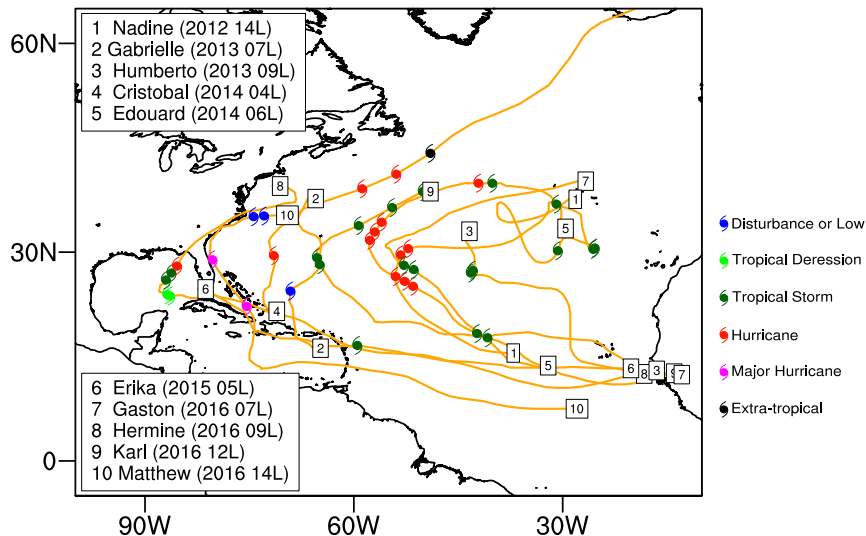


FIG. 1. Locations and intensities of available cases with GH dropwindsonde observations. TCs 1–5 were flown by NASA HS3, while TCs 6–10 were flown by NOAA SHOUT.

longer time window, such as 24 h, may include influences from interactions with environment (e.g., dry air and shear). With this estimate, when the intensification rate during this 12-h period is $\geq 20 \text{ kt (24 h)}^{-1}$, the case is labeled as intensifying; when the rate is between -20 and $20 \text{ kt (24 h)}^{-1}$, the case is labeled as steady state; and when the rate is $\leq -20 \text{ kt (24 h)}^{-1}$, the case is classified as weakening. This intensification-rate classification, shown in Fig. 2, depicts the frequency distribution of the cases based on intensity. Most of the tropical depressions or weak tropical storms ($< 50\text{-kt}$ intensity) are steady state. Intensifying cases are strong tropical storms ($\geq 50\text{-kt}$ intensity) or hurricanes. The weakening TCs are mostly hurricanes ($> 60\text{-kt}$ intensity). In total, there are seven intensifying cases, 27 cases at steady state, and six weakening cases.

b. Data assimilation system: HEDAS

The Hurricane Ensemble Data Assimilation System (HEDAS) is a research DA system that consists of an ensemble square root Kalman filter (Whitaker and Hamill 2002) and the Hurricane Weather Research and Forecasting (HWRF) Model (Gopalakrishnan et al. 2013). It assimilates high-resolution observations at the vortex scale in either an Earth-relative or storm-relative framework (Aksoy 2013). Christophersen et al. (2017) demonstrated that assimilating GH dropsondes using HEDAS results in a positive impact on the prediction of TC track, intensity, and structure in two cases of Hurricane Edouard (2014). The present study uses the same HEDAS configuration as in the previous experiments. Deterministic forecasts from final ensemble-mean analyses are carried out using the same HWRF configuration as in DA.

c. Experiment setup

As discussed in Christophersen et al. (2017), the presence of crewed reconnaissance aircraft observations, especially tail Doppler radar (TDR) radial wind velocity superobservations (superobs; Gamache 2005), is critical for representing the complex TC vortex structure. Hence, the first set of experiments is run by withdrawing observations from all crewed aircraft (P-3, G-IV, and U.S. Air Force), including high-density flight-level observations, Stepped Frequency Microwave Radiometer (SFMR; Uhlhorn et al. 2007), TDR, and dropsondes. A set of parallel experiments with and without GH dropsondes is then carried out to assess the impact of GH dropsondes without the influence of crewed aircraft observations. Other datasets assimilated when available in this scenario are the Atmospheric Infrared Sounder (AIRS) cloud-cleared retrievals (Susskind et al. 2003), Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) global positioning system (GPS) radio occultation retrieved profiles (Kuo et al. 2004), nearby rawinsondes, and flight-level data from the Aircraft Communication, Addressing, and Reporting System (ACARS).

The second set of experiments is run with observations from all crewed aircraft included to evaluate the impact of GH dropsondes when data from crewed aircraft are present. Table 1 lists the number of cases per TC when crewed reconnaissance and/or surveillance aircraft are available.

For each case, the first 30 out of 80 members of NOAA's ensemble-based Global Forecast System analyses (Hamill et al. 2011) are used as the initial and boundary conditions for the HWRF ensemble during the 4-h spinup and

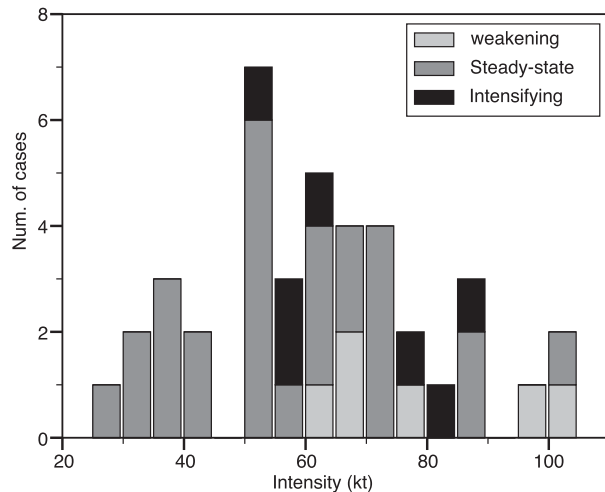


FIG. 2. Stacked bar plot of the intensity frequency for all cases in Table 1 with stratification according to 24-h intensity change.

subsequent 4-h DA cycling (± 2 h of the synoptic time). The DA is only performed on the vortex-following $10^\circ \times 10^\circ$ inner domain with 3-km grid spacing that is embedded in the 9-km-resolution outer domain with two-way interaction. All Earth-relative observations within a 6-h time window are assimilated in a storm-relative framework (Aksoy 2013) in 30-min cycles (a total of nine cycles). The last analysis after the nine cycles is called the “final” analysis in this study. Note that the number of cycles (nine in this case) and the length of the DA cycling window (4 h in this case) can be flexibly configured independent of the 6-h observation window because storm-relative observations are randomly redistributed in time, as in Aksoy (2013). Five-day deterministic forecasts for each case are initialized from the ensemble mean of the final analyses. More detailed description of the setup of the DA and model for the experiments can be found in Christophersen et al.’s (2017) Table 1.

3. GH dropsondes overall cycling performance

The GH dropsonde data are taken from the traditional TEMP-DROP format that provides data at mandatory and significant levels, where the location and time at each level is estimated from high-resolution launch and splash positions, times, and measured wind velocities; hence, the horizontal and vertical drifts of the dropsonde as it falls are accounted for (Aberson et al. 2017). Data quality-control and thinning details can be found in Hock and Franklin (1999), and they are consistent among various field campaigns. Observation errors follow the specifications in Aksoy et al. (2013). It is noted that the relative humidity observations from GH dropsondes before the 2016 hurricane season are not corrected for a recently discovered

dry bias (Vömel et al. 2016); these observations are mostly above 400 hPa and are assigned large observation errors (0.5 g kg^{-1}) so that they would not have a significant influence on the analyses (Christophersen et al. 2017).

The overall performance of GH dropsondes is assessed for cases from all 10 TCs. Figure 3 shows the contoured frequency by altitude diagram (CFAD; Yuter and Houze 1995) of innovations for temperature T , specific humidity q , zonal wind speed u , and meridional wind speed v at observation locations for the first-cycle prior (forecast) and last-cycle posterior (analysis) during DA. Slightly more observations of wind and temperature are assimilated than moisture, likely because relatively stronger gradients of wind and temperature in the upper atmosphere (≥ 200 hPa) result in more significant levels in TEMP-DROP messages. Compared to T and wind, q presents a more even distribution vertically.

The salient difference between the forecast and analysis innovation frequencies for all observation types is that the tails of the distributions are largely eliminated by the end of cycling. The distributions are also less skewed in the analyses, compared to the forecasts. Finally, the analyses have smaller values of kurtosis, indicating their distributions resemble the Gaussian distribution better than those of the forecasts. These statistics indicate that the data are less biased and more Gaussian after DA and, hence, are more compatible with the assumptions of the DA system used here.

Specifically, T exhibits an initially negative bias below 700 hPa (possibly suggesting that GFS background fields are warmer in the lower troposphere than observations, or an HWRF model bias from the spinup) and a secondary peak at around 500 hPa. After DA, the mode is at 700–900 hPa, although there is still a slight negative bias in the planetary boundary layer (PBL; Fig. 3a). For q , the persistent higher frequencies at negative innovations (Fig. 3b) indicate that a dry bias possibly introduced by the initial GFS background is not entirely eliminated by DA. There is a slight negative bias for u in the lower troposphere initially, which is mostly removed after DA (Fig. 3c). In terms of v , distributions exhibit positive bias initially in the lower and midtroposphere, likely due to underestimation of intensity in the low-resolution GFS background fields. After DA, this positive bias is corrected (Fig. 3d).

4. Overall performance without assimilation of crewed aircraft observations

a. Analyses

The overall analysis quality after the assimilation of GH dropsondes in the absence of crewed aircraft observations

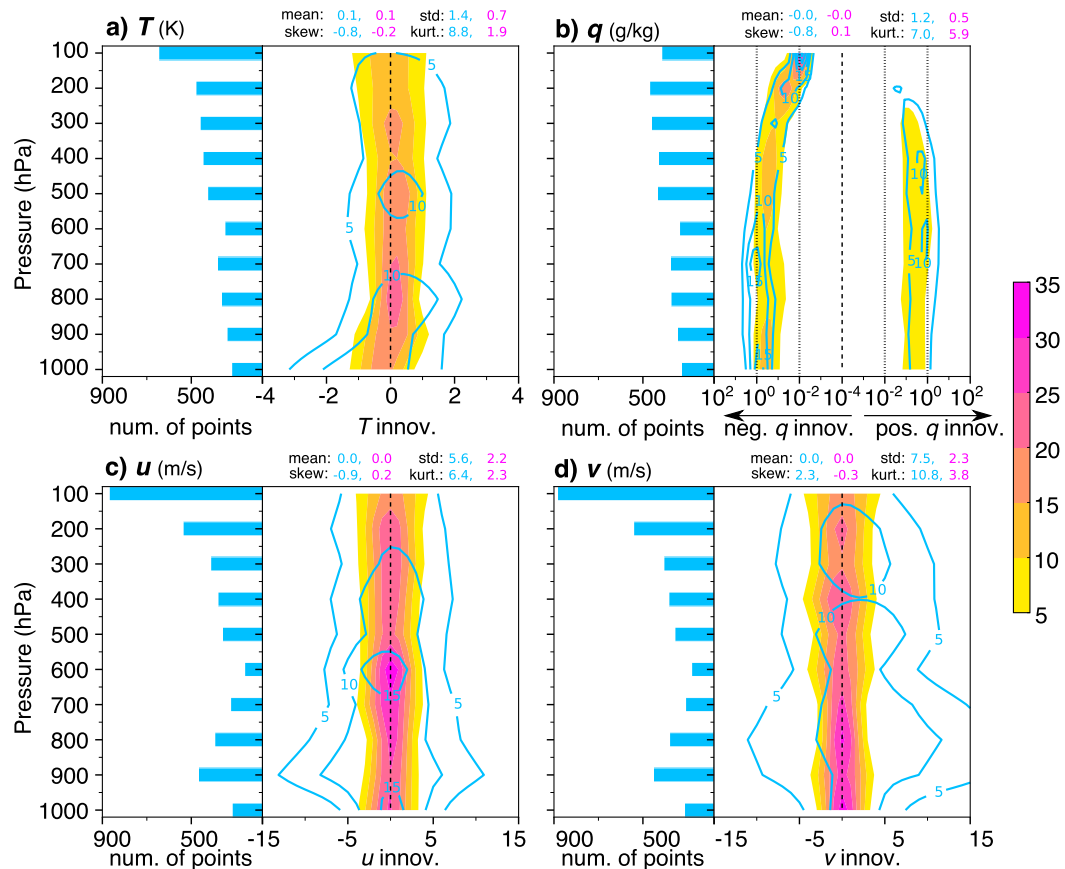


FIG. 3. CFADs for the innovations of (a) temperature T , (b) specific humidity q , (c) zonal wind speed u , and (d) meridional wind speed v for the first DA cycle's forecast (blue contours) and the last DA cycle's analysis (shaded). Innovations of q are shown on a logarithmic scale. Data distribution of T , q , u , and v as a function of pressure for the first cycle's forecast is also shown on the left for each panel.

is assessed by verifying TC position, intensity, minimum sea level pressure (MSLP), and the wind–pressure relationship in the final mean analyses against the best track (Fig. 4). The best track estimates are provided by NHC as standard b-deck files and follow Jarvinen et al. (1984). The centers in the model are tracked by the MSLP. In general, the mean position, intensity, and MSLP improvements from assimilating GH dropsondes for all of the cases is about 16%, 15%, and 20%, respectively (see Table 2). When GH dropsondes are assimilated, the intensity and MSLP exhibit better linear fits to the best track (Figs. 4b,c), although a general bias persists, indicating underestimation of intensity, similar to findings by Aksoy et al. (2013) and Aberson et al. (2015). The mean absolute distance between the intensity linear regression lines for cases with and without GH dropsondes is 5.7 and 6.5 kt, respectively; for MSLP, the mean absolute distance is 10.4 and 13.4 hPa, respectively. Also, the improvement in analyzed intensity for stronger cases (intensity ≥ 65 kt) is noteworthy. The wind–pressure relationship for the experiments with GH

dropsondes also shows closer linear fit than the one without GH dropsondes when compared to the best track fitted line (Fig. 4d). The mean absolute distance in this scenario for cases with and without GH dropsondes is 6.0 and 7.2 hPa, respectively.

The frequency distributions of the position and intensity errors for all cases in the final analyses are also shown (Fig. 5). The distributions indicate that the mean errors are reduced when GH dropsondes are assimilated. There is a cluster of cases with smaller position errors when assimilating GH dropsondes. With the assimilation of GH dropsondes, the standard deviations for TC intensity and MSLP absolute errors also are reduced (Table 2). This shows that analyses with GH dropsondes assimilated exhibit better intensity error consistency (smaller standard deviations) among cases than analyses without GH dropsondes. However, the standard deviation of the absolute position errors with GH dropsondes is larger than that without GH dropsondes. This is most likely due to a combination of a few outliers

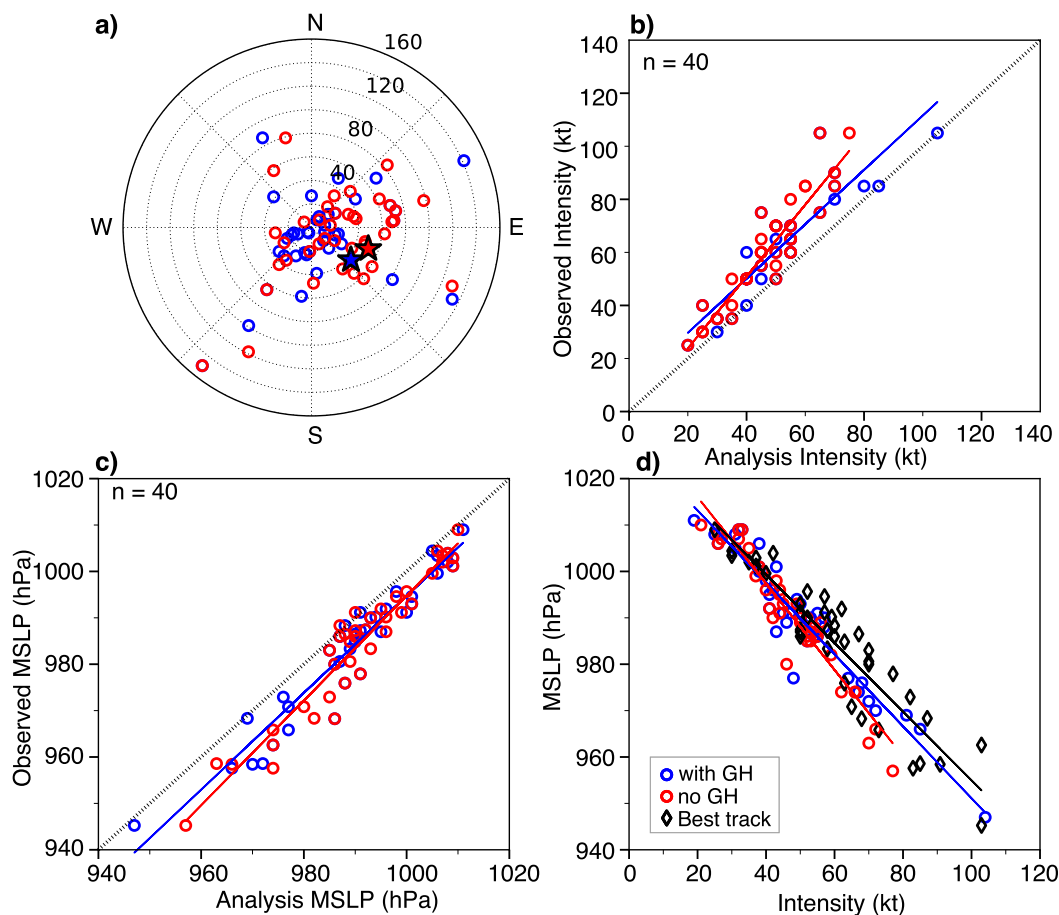


FIG. 4. Scatterplots showing results from experiments with (blue circles) and without (red circles) GH dropsonde data for (a) position error [the stars indicate mean position errors for the experiments with (blue) and without (red) GH dropsonde data], (b) intensity, and (c) MSLP in the final analyses, compared to the best track. (d) The wind–pressure relationship. The blue, red, and black lines in (b)–(d) represent linear regression fits. The dashed light black lines in (b),(c) denote the 1:1 lines.

and a smaller mean, which renders larger deviations from the mean, compared to the scenario of a greater mean with similar outliers (Fig. 5a). By examining the median absolute deviation (MAD; a more robust measurement of the variability of the sample; Wilks 1995) of the absolute position errors, it is clear that overall position errors are still reduced with the assimilation of GH dropsondes (Table 2). Visual inspection of Fig. 5a also confirms that the resulting frequency distribution becomes much narrower around the mode when GH dropsondes are assimilated, indicating

that the position error variability is smaller for most cases. However, a Wilcoxon rank-sum test (Wilks 1995) shows that the two distributions for position, intensity, or MSLP are not statistically significant at the 95% significance level.

It is noted that TCs moving at a fast speed (>35 kt) are the outliers we see in Fig. 5, either due to large position errors or large intensity/MSLP errors. For example, Karl (2016) on 25 September and Cristobal (2014) on 29 August moved northeastward at a speed of >40 kt and had large position errors ($\sim >100$ km) regardless of

TABLE 2. The mean, standard deviation (Std), and MAD of position errors, MSLP errors, and intensity errors at the final analyses for the experiments with and without GH dropwindsondes.

	Position (km)			Intensity (kt)			MSLP (hPa)		
	Mean	Std	MAD	Mean	Std	MAD	Mean	Std	MAD
With GH	43.3	37.0	14.9	10.6	7.9	2.7	−5.7	4.2	4.5
Without GH	51.5	31.6	19.2	13.3	8.5	2.7	−6.5	4.5	6.0

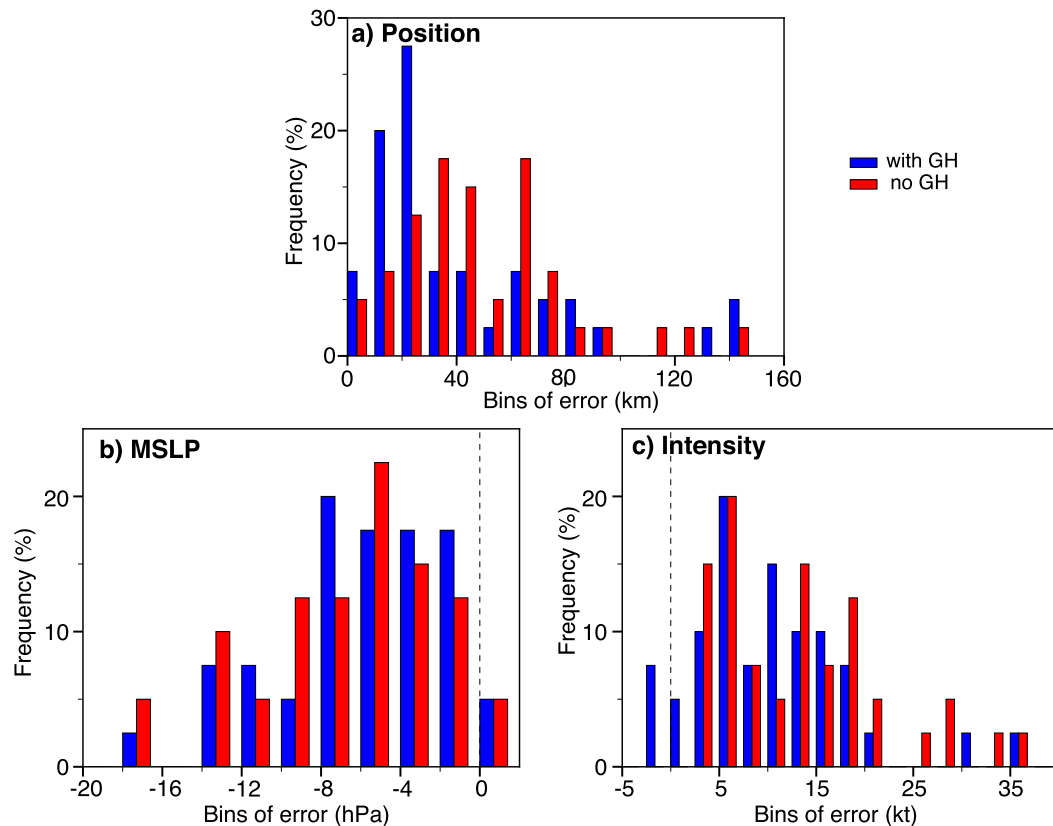


FIG. 5. Frequency distributions of (a) position, (b) MSLP, and (c) intensity errors in the final analyses, compared to the best track for experiments with and without GH dropsondes assimilated.

whether GH dropsondes were assimilated. Other outliers with large intensity bias (>30 kt) are cases where GH sampled only the outer environment with no additional inner-core observations to correctly represent the vortex structure during DA, such as Matthew (2016) at 1800 UTC 5 October and Edouard (2014) at 1800 UTC 14 September.

The structures of TCs in the final analyses are evaluated in Fig. 6 by comparing integrated kinetic energy (IKE; Powell and Reinhold 2007), quadrant-mean maximum radius of 34-kt winds (R34; average of the maximum 34-kt wind speed radii reported in the four storm quadrants; any quadrants that do not contain a maximum 34-kt wind speed radii are excluded from the averaging), and radius of maximum wind speed (RMW). The calculations of IKE and R34 follow the same procedure for both the model analyses and the best track estimates, although the observational quantities have larger uncertainties than the model due to observation sparsity. IKE is a quantity that combines TC intensity and size (Powell and Reinhold 2007). The calculation of IKE follows Aberson et al. (2015). In addition, sample correlations among various size and

intensity parameters in Fig. 7 are calculated for the cases examined, separately for analyses and the best track, where all radii (R34, R50, and R64) are again the averages of individual values reported in the four quadrants.

Comparing Fig. 6a to Figs. 4b and 4c, TC structure exhibits more variability than intensity. Overall, cases with GH dropsondes present a higher variance explained for the linear fits than those cases without. The HEDAS analyses underestimate IKE, compared to the best track, which is likely a combination of underestimation of intensity (Fig. 4b), underestimation of R34, especially for smaller TCs (Fig. 6b), and overestimation of the RMW (Fig. 6c). However, variations of IKE are nonlinearly dependent on the TC size and intensity (Powell and Reinhold 2007; Aberson et al. 2015); thus, the factors that contribute to IKE errors are difficult to diagnose. Therefore, we further examine the correlations between IKE and various TC structure metrics (Fig. 7). For the experiments with or without GH dropsondes, IKE is strongly correlated with intensity, R34, and RMW (Figs. 7a,b), while for the best track, IKE is strongly correlated with R34 and R50 (Fig. 7c). Nonetheless, the correlation

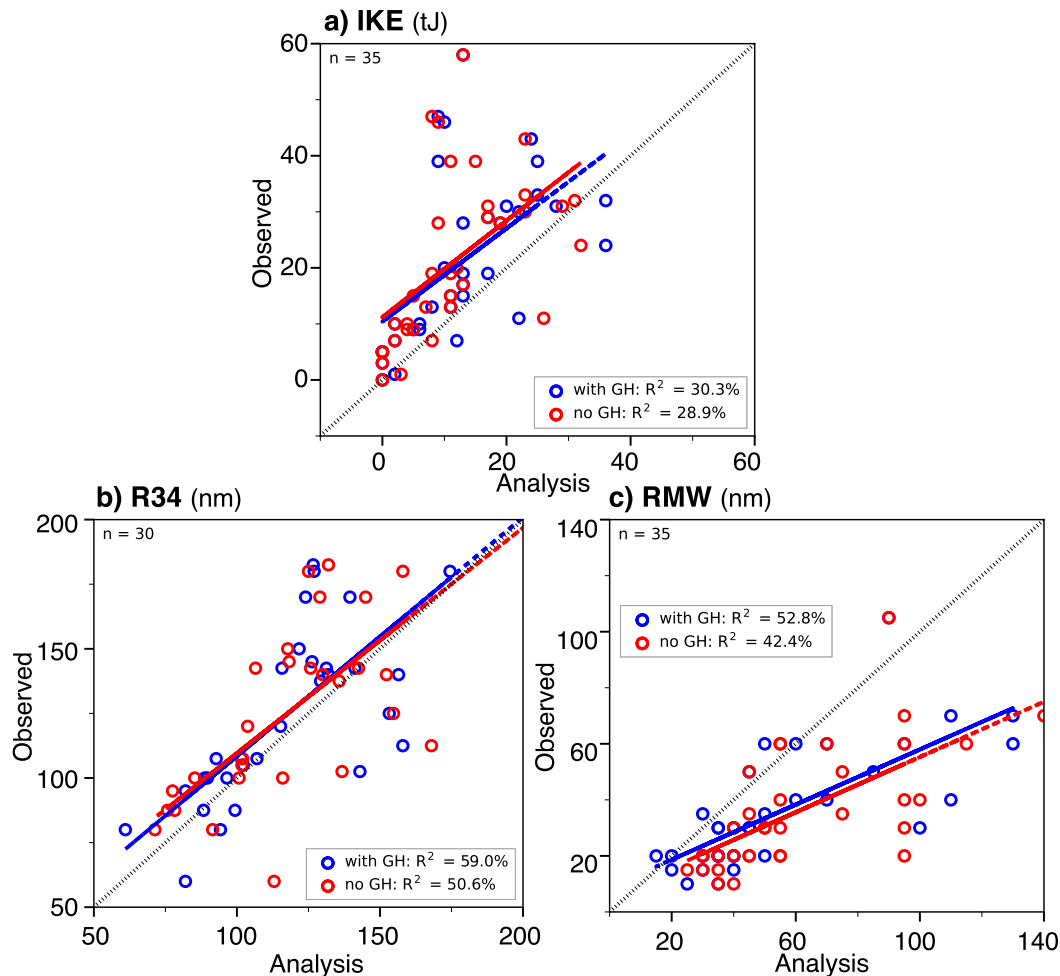


FIG. 6. Scatterplots of (a) IKE, (b) R34, and (c) RMW between analyses and the best track for cases with tropical storm or stronger intensity.

between IKE and R34 is consistently high in both the analyses and the best track. To further support this finding, the correlation between IKE and the quadrant-maximum radius of 34-kt winds radii (the maximum of 34-kt winds radii maximum extent reported in the four quadrants) is also high, compared to that of IKE and R34 (not shown). Stepwise regression analysis (Wilks 1995) between IKE and various TC metrics shows that R34 is the dominant predictor in both the analysis and the best track in terms of either statistical significance or adjusted variance explained. Hence, the underestimation of R34 is hypothesized to be the dominant source of error in IKE.

We note that both the analyses and the best track show small correlations between TC intensity and size (e.g., correlation between intensity and R34, R50, or R64). We further note that the relatively small sample size limits our ability to make generalized conclusions

beyond the current study, even though some of these correlations are statistically significant.

b. Deterministic forecasts

The overall errors and relative skill of the deterministic forecasts initialized from the final ensemble-mean analyses are shown in Fig. 8. In general, we see that the assimilation of GH dropsondes generally results in improvements to both track and intensity forecasts by up to 20%. Specifically, there is a slight improvement to track forecasts up to 84 h when GH dropsondes are assimilated, and most of these improvements are at least 90% statistically significant. The relative skill of the track forecasts is maximized at 12 h ($\sim 20\%$) and then gradually decreases to nearly neutral by 84 h. The track errors are mostly in speed (along track) during the first 48 h and cross track beyond that. Assimilation of GH dropsondes does not seem to modify the nature of the track error

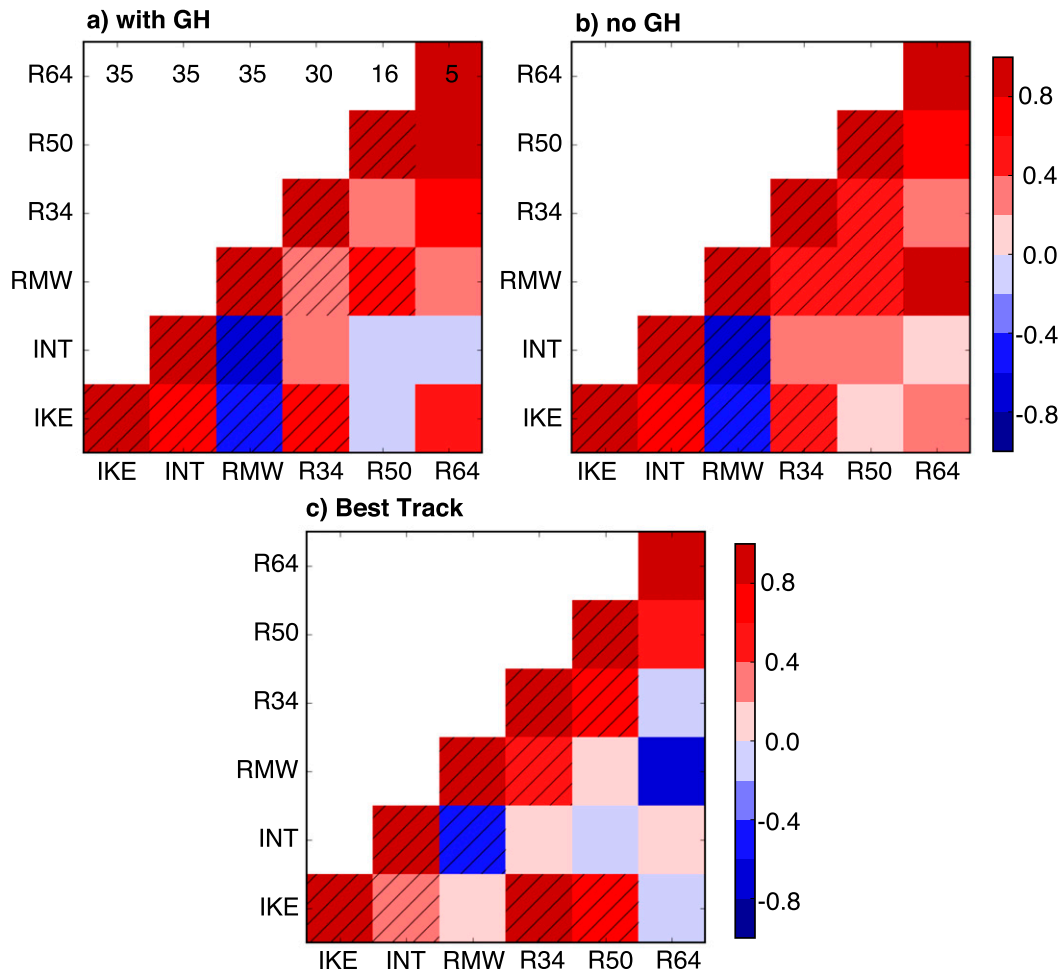


FIG. 7. Correlation matrix for IKE, intensity (INT), RMW, R34, R50, and R64 (see text for definition) for experiments (a) with GH dropsondes and (b) without GH dropsondes, and (c) the best track. The number of cases is also denoted in the row of R64 in (a). Hatched areas indicate correlations that are at least 95% statistically significant.

distributions (not shown). Meanwhile, the MSLP forecasts with the assimilation of GH dropsondes present mostly consistent superior skill ($\sim 10\%$) for up to 108-h lead time.

5. Dependence of performance on intensity change

We further address the analysis and forecast performance as a function of TC intensity change as classified in section 2a. Note that the performance calculations presented in this section continue to use cases without assimilation of crewed aircraft observations. To reduce sample size deficiency in the three categories (see section 2a), the intensifying and weakening cases are combined into one group, labeled as nonsteady state (non-SS). This is mainly justified by the fact that the assimilation of GH dropsondes results in comparable forecast skill in the intensifying and weakening cases,

although these cases exhibited various TC environments, and it was difficult to observe commonalities in TC structure due to limited sample size. To further eliminate the potential influence of TC intensity on the statistical outcomes, weak tropical storms and tropical depressions (TC intensity less than 50 kt) are not included in steady-state cases. Sample sizes thus obtained in the intensification-based stratification are 19 for SS cases and 13 for non-SS cases, both in the intensity range of 50–105 kt.

a. Analyses

We first examine mean position and intensity errors, compared to the best track estimates (linearly interpolated at 30-min cycling frequency) during DA cycles. Figure 9 shows mean position and intensity errors as a function of cycle time during the 4-h DA window (± 2 h centered on the synoptic time), calculated separately for

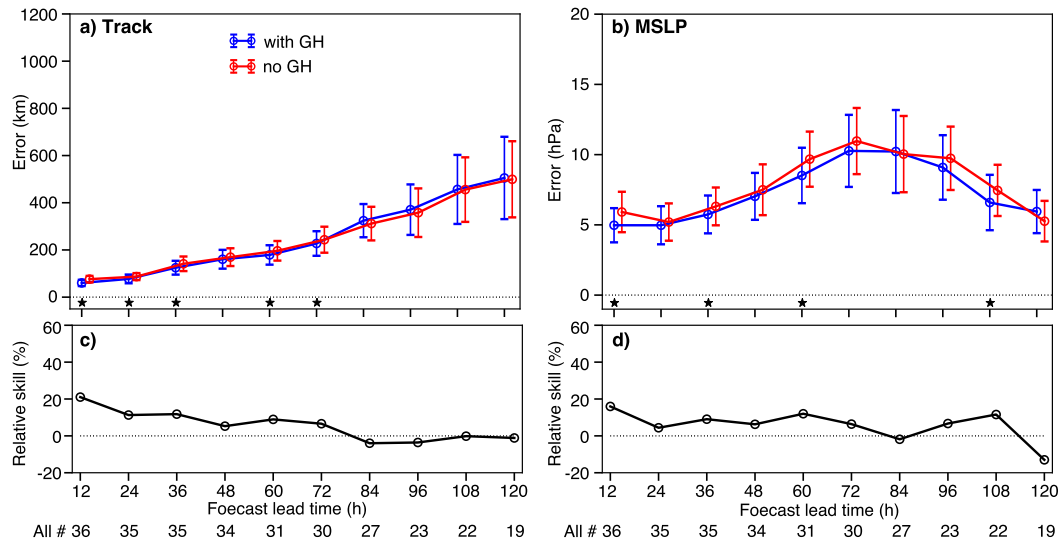


FIG. 8. Average (a) track and (b) MSLP forecast errors and (c),(d) their corresponding relative skill (% improvement) for cases with and without dropsondes assimilated. The 95% confidence intervals for the errors at each forecast lead time are shown in error bars. The forecast errors at a given forecast lead time that are at least 90% statistically significant are indicated in stars in (a),(b). The errors and confidence intervals for cases without dropsondes are slightly displaced in x axis to avoid overlapping with the ones with dropsondes.

SS and non-SS cases. The sawtooth pattern indicates the change in error (from forecast to analysis) during DA at a particular cycle (vertical lines) versus the change in error (from analysis to forecast) during model integration between two cycles (oblique lines). The sawtooth pattern is typical in a cycling DA system (e.g., Aksoy et al. 2009, 2012) and ideally indicates the balance between the decrease in error from assimilating observations at DA cycles and the growth of forecast error between DA cycles. In general, the SS cases present larger position errors and confidence intervals than the non-SS cases. Furthermore, with GH dropsondes assimilated, a greater reduction in position errors during DA cycles is achieved in non-SS cases than SS cases. It is unclear why SS cases have larger position errors or larger spread of position errors than non-SS cases, regardless of whether or not the GH dropsondes are assimilated. However, as suggested from the larger average 12-h track error in SS versus non-SS cases (172.8 vs 80.4 km, respectively) in the climatology and persistence (CLIPER; Aberson and Sampson 2003) statistical model, the track in the SS cases on average is expected to be more difficult to predict than the non-SS cases against climatology. We therefore speculate that the smaller analysis position errors in the non-SS cases are a general result of the track of these cases being climatologically easier to predict than the SS cases (assuming 12-h linear error growth within the CLIPER model).

For intensity and MSLP (Figs. 9b,c), SS cases in general exhibit smaller errors than non-SS cases during DA,

which is the opposite of what was observed for position errors (Fig. 9a). Furthermore, the non-SS cases without GH dropsondes exhibit noticeable increases in error during cycling. This is likely due to HWRF without the assimilation of GH dropsondes not being able to capture intensity change during cycling. Nevertheless, including GH dropsondes in DA better captures changes in TC evolution, resulting in a larger reduction of MSLP and intensity errors by the end of cycling relative to the one without-GH dropsondes. Meanwhile, the assimilation of GH dropsondes shows a relatively small impact on the mean intensity of the SS cases. GH dropsondes also lead to a gradual reduction in IKE errors during 4-h DA cycling for non-SS cases, although steady IKE error growth occurs regardless of whether GH dropsondes are assimilated (Fig. 9d). Little impact on IKE is seen for SS cases. IKE errors exhibit larger errors in the SS cases than the non-SS cases, similarly as the position errors. However, IKE evolves similarly as intensity, especially in the non-SS cases without GH, where an unbounded error growth is seen in both intensity and IKE.

It is known that dynamical processes involved in weakening and intensification of TCs are rather different. However, despite such differences, it appears that sampling the TC with high-altitude dropsondes from GH helps reduce the errors that result from these processes. Overall, it is evident that GH dropsondes demonstrate a larger impact for TCs that undergo relatively large intensity change during DA. As a result, we see larger relative skill in the final analyses for the non-SS

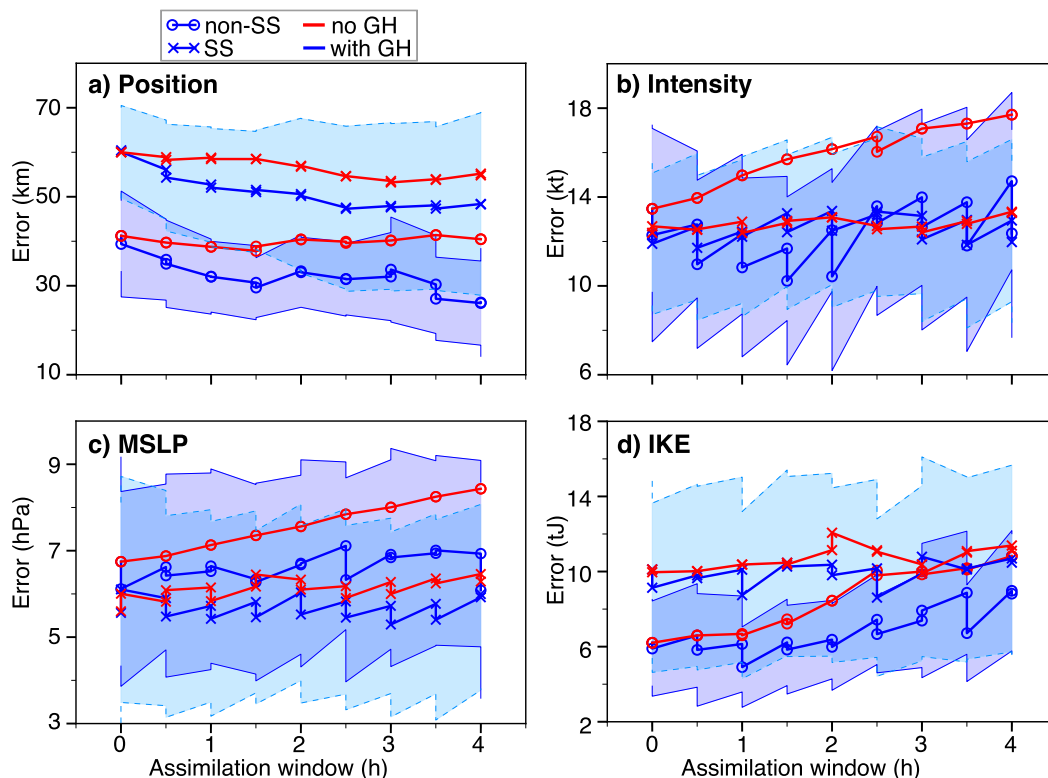


FIG. 9. Average TC (a) position, (b) intensity, (c) MSLP, and (d) IKE absolute errors, as compared to the best track during the 4-h DA window for SS and non-SS strong cases (≥ 50 kt). The best track is interpolated to the 30-min DA cycle times. Shaded areas are the 95% confidence intervals for SS (light blue shading with dashed edges) and non-SS cases (light purple shading with solid edges) in the experiment with GH assimilated.

cases than the SS cases (Fig. 10). The relative skill from assimilating GH dropsondes for non-SS cases ranges between 25% and 35% for either the position or intensity improvement, while it is only ~5%–10% for SS cases.

The impact of GH dropsondes on the azimuthal wavenumber-0 structure of the SS and non-SS cases is shown in Fig. 11. The larger impact from assimilating GH dropsondes occurs in non-SS cases, where visible modulations of azimuthally averaged tangential wind speed, radial wind speed, and equivalent potential temperature anomaly θ'_e are evident. The assimilation of GH dropsondes appears to intensify the tangential and radial wind speeds, as well as increase the θ'_e , while also resulting in a smaller vortex (smaller RMW) on average. We notice that there is little change in the wavenumber-0 structure for SS cases (Figs. 11d–f). Overall, the impact on wavenumber-0 structure is consistent with the overall analysis performance during DA and in the final analyses. These larger modulations of structure along with smaller position errors likely contribute to the superior skill in the final analyses for non-SS cases over SS cases (Fig. 10).

b. Deterministic forecasts

Position and MSLP forecasts are verified against the best track at an interval of 12 h (Fig. 12). Forecast errors for both track and MSLP beyond 24-h lead time are

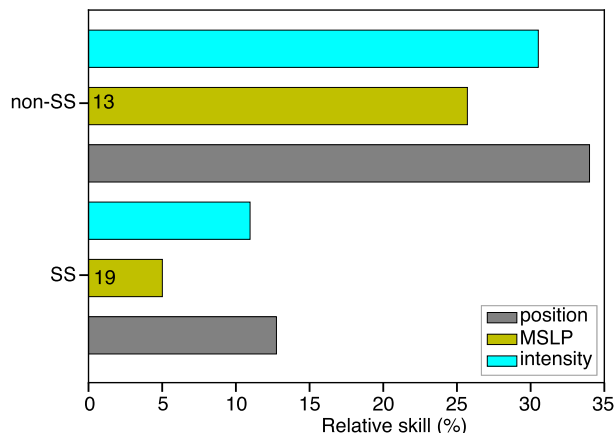


FIG. 10. Relative skill (% improvement) for TC position, MSLP, and intensity in the final analyses from experiments with GH dropsondes vs without for SS and non-SS TCs with intensities ≥ 50 kt. Number of cases is also indicated on the MSLP bars.

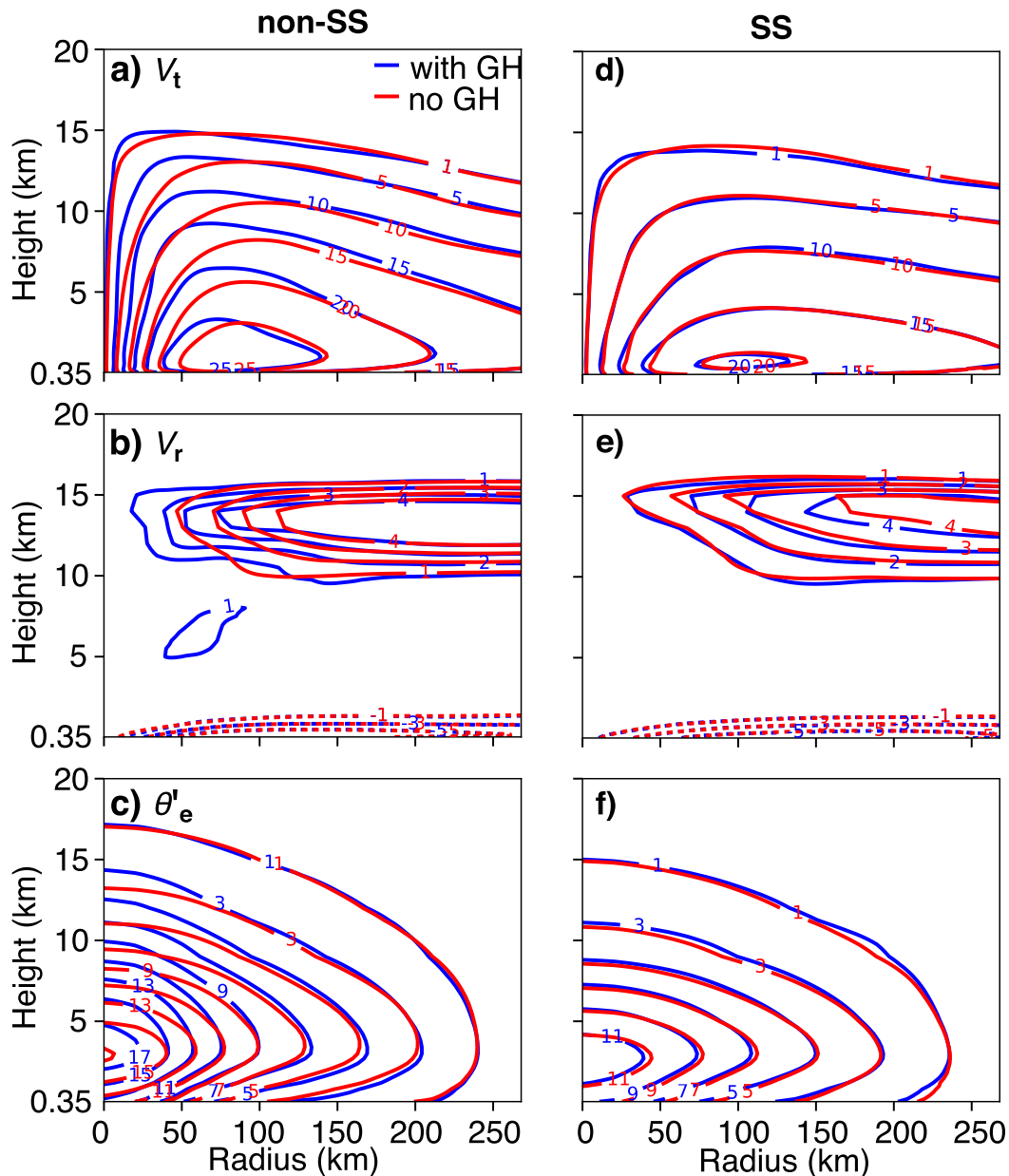


FIG. 11. Composite azimuthal averages of (a),(d) tangential wind speed; (b),(e) radial wind speed; and (c),(f) equivalent potential temperature anomaly θ'_e for (a)–(c) non-SS and (d)–(f) SS TCs.

smaller for non-SS cases than SS cases. Furthermore, GH dropsondes have a positive impact on track forecasts for non-SS cases nearly throughout the entire 5-day forecast. In particular, the relative improvement remains above 15% for lead times up to 72 h. The slight degradation at the longer lead times of 84–84 h is attributed to two Hurricane Hermine (2016) cases, where a slightly slower TC motion with GH dropsondes caused the track at later forecast times (84–120 h) to be very different because of complex interactions with an upper-level approaching trough and

the TC undergoing extratropical transition (ET) while the TC maintained tropical-storm-force winds. Forecasting of TC track in such complex scenarios that involve ET remains a challenge in operations (Jones et al. 2003; Evans et al. 2017). However, the 84–96-h degradation is shown not to be statistically significant (Fig. 12a).

For the MSLP forecasts of non-SS cases, a small positive impact is seen in the first 24 h, followed by a more consistent positive impact ($\sim 20\%$ improvement) beyond 48-h lead time. The degradation of the forecasts between 24 and 48 h

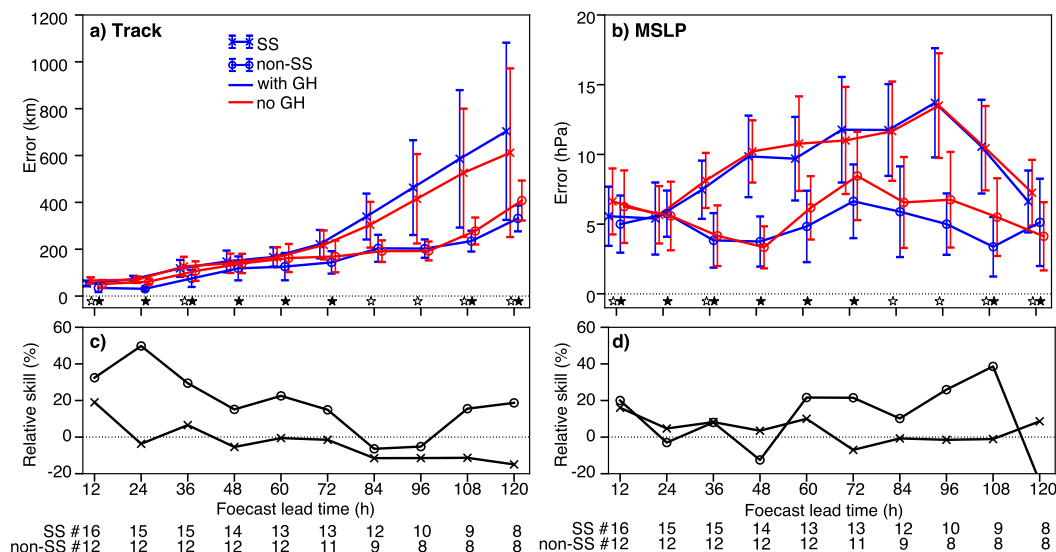


FIG. 12. Average (a) track and (b) MSLP absolute forecast errors and (c),(d) their corresponding relative skill (% improvement) for strong (with intensities ≥ 50 kt) non-SS and SS cases with and without GH dropsondes assimilated. The 95% confidence intervals for the errors at each forecast lead time are shown in error bars. The forecast errors at a given forecast lead time that are at least 90% statistically significant are indicated in stars in (a), (b), where the open stars are for SS cases and filled stars are for non-SS cases. The displacement in x axis at each forecast lead time for the errors and confidence intervals of the cases without dropsondes is only applied to avoid overlapping with the ones with dropsondes.

is associated with cases that undergo eyewall replacement cycles (ERCs), such as Edouard (2014) on 16 September and Gaston (2016) on 29 August. For these two cases, the GH dropsondes mainly sampled the outer environment of the TCs and provided limited coverage of the inner core. Therefore, the two cases produce analyses with large intensity errors (>20 -kt errors, compared to the best track) at the initial time even with the assimilation of the GH dropsondes. It is also possible that the important dynamical processes (e.g., ERC) involved may not be well represented by the measurements, which could contribute to the forecast degradation for the two cases. Furthermore, intensity prediction during ERC continues to be a challenging task in research and operational practice. The degradation seen at 120 h is mainly from the Gaston (2016) case at 1200 UTC 27 August, where a faster forward motion of the storm is seen in the experiment with GH dropsondes assimilated after the TC turns east-northeastward beyond 36-h forecast lead time, resulting in an intensity forecast degradation after that time. Nevertheless, it is likely that such variability of error with forecast lead time is also due to limited sample size. We further hypothesize that the storm-relative preprocessing in HEDAS, which assumes the TC structure is nearly stationary during the 6-h observation period (Aksoy 2013), may be limiting the performance of GH dropsondes in fast-moving cases, such as TCs that undergo ET, especially when the number of center fixes available is limited to interpolate observation locations in between. Optimization of

assimilating dropsondes in these cases needs to be examined and is beyond the scope of the current study.

6. The impact of GH and crewed aircraft data on TC forecasts

The impact of GH dropsondes on TC forecasts is also examined for cases with crewed aircraft data available at the analysis time. Here, the denial experiments are specifically based on the availability of TDR data, although these data are almost always available along with flight-level, SFMR, and dropsonde data from the P-3 or G-IV aircraft in these cases.

The overall forecast performance for all cases with and without crewed aircraft data is shown in Fig. 13. With crewed aircraft data, assimilating GH dropsondes leads to mostly superior track forecasts throughout the 5-day period. In fact, relative skill improvements to the track forecasts are $\sim 10\%$ – 20% (Fig. 13a). The positive impacts on the track forecast are also at least 90% statistically significant beyond the 72-h lead time. However, in the case of intensity and MSLP, no obvious benefit is seen from assimilating GH dropsondes at most forecast lead times (Fig. 13c). The degradations of the intensity forecasts, however, are not statistically significant at most lead times, except at 12 and 84 h, and are thus subject to large variability among cases. For instance, the intensity degradations in two Edouard (2014)

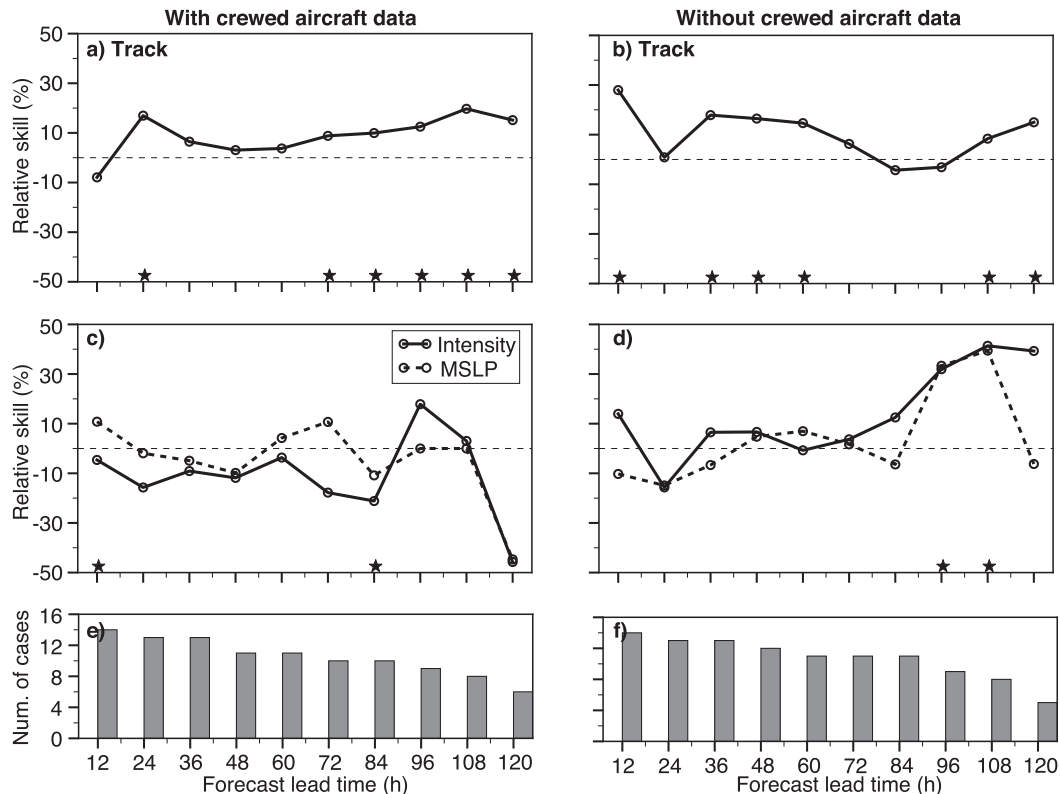


FIG. 13. Relative skill of assimilating GH dropsondes over not assimilating GH dropsondes for (a),(b) track, and (c),(d) intensity and MSLP averaged for all cases with crewed aircraft data available. The relative skill at a given forecast lead time that is at least 90% statistically significant is indicated in stars in (a)–(d), where in (c),(d) the significance is for MSLP. The baseline experiment in the left panel includes crewed aircraft data, while the baseline one in the right panel does not. (e),(f) The number of cases at each forecast lead time for the two sets of experiments.

cases are related to the timing of the secondary eyewall replacement (SEF) and ERC during forecasts, while the degradations from Hermine (2016), Karl (2016), and Cristobal (2016) are largely attributed to errors in storm speed that result in the analyzed TCs to become out of phase in how they interact with their environments at various forecast lead times. Without crewed aircraft data, GH dropsondes lead to mostly positive track forecast skill throughout the forecasts (Fig. 13b) and positive intensity and MSLP forecast skill beyond 1-day lead time (Fig. 13d). The impact in the absence of crewed aircraft data is greatest at long lead times (i.e., 96–120 h) with values of over 30%. The ~30% improvements at 96- and 108-h lead times are at least 90% statistically significant. Further diagnosis indicates that strong TCs (tropical storms or hurricanes at initial times) dominate these track and intensity improvements in the absence of crewed aircraft data. Regardless, the overall results demonstrate that GH dropsondes are capable of providing track and intensity forecast improvements, especially in the absence of crewed aircraft data.

7. Track forecast examples

Two cases characterized by large track improvements with the assimilation of GH dropsondes are discussed. In the first case, Tropical Storm Edouard (2014) at 0600 UTC 12 September was located over 1000 n mi (1852 km) east of the Lesser Antilles. Edouard at this location could only be reached by the UAS, such as GH, with >2-h on-station time. A lawnmower pattern was used to sample the TC. The assimilation of GH dropsondes resulted in small modifications on the initial-time vortex structure (e.g., tangential and radial winds) and thermodynamic environment near the TC. The TC in both forecasts was then steered by the deep-layer subtropical ridge in the central and eastern Atlantic Ocean. The improved track forecast in the first 36 h is likely due to the assimilation of GH dropsondes. By 60 h, slightly more westward upper-level flow was seen in the forecast with GH dropsondes (Fig. 14b). This westward shift of upper-level steering continued until 108 h, when the TC made its northeastward turn. Further examination at 60 h reveals that the ridge to the northeast of the

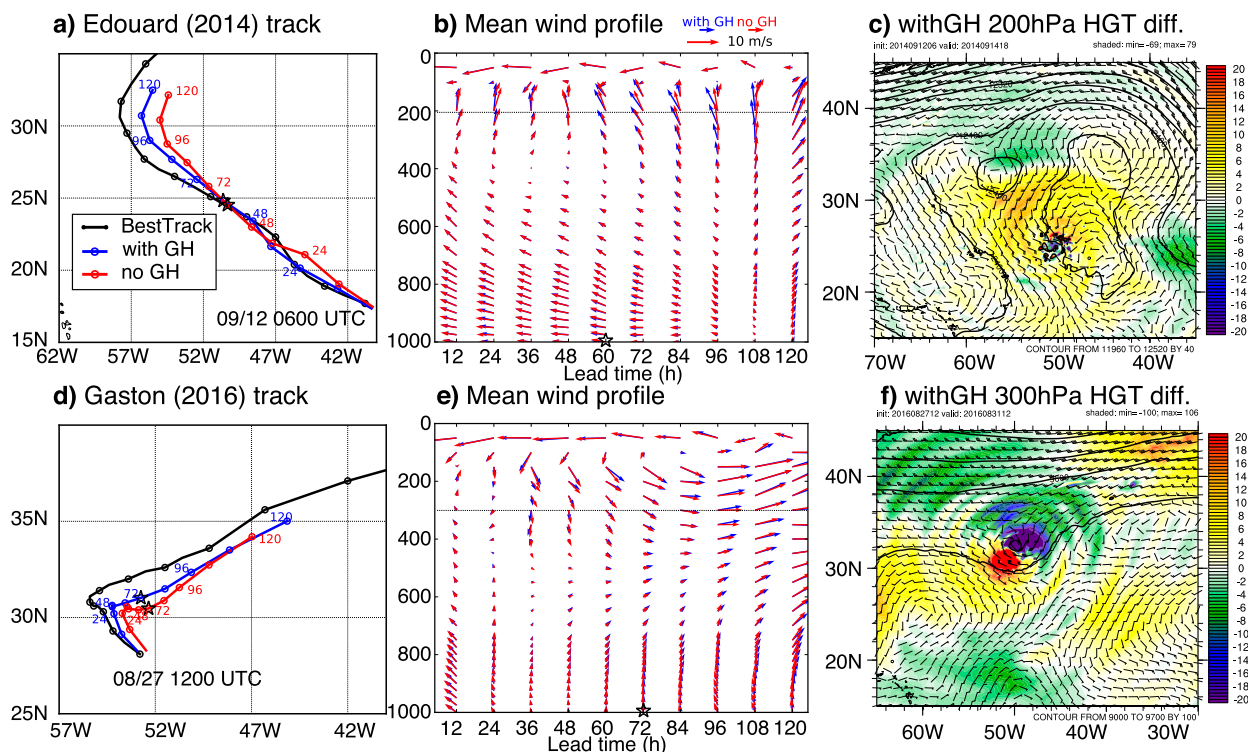


FIG. 14. (a) Five-day forecast tracks for Edouard (2014) initiated at 0600 UTC 12 Sep; (b) mean wind profiles averaged within 500 km of the TC center; and (c) 200-hPa geopotential height (gpm; contours) and wind vectors (m s^{-1}) from the experiment with GH dropsondes and the differences of geopotential height (with GH minus no GH; shaded) at lead time 60 h. (d)–(f) As in (a)–(c), but for the case of Gaston (2016) initiated at 1200 UTC 27 Aug. The geopotential height, wind vectors, and differences in (f) are at 300 hPa.

TC was stronger in the forecast with GH, thus allowing more northwestward steering (Fig. 14c) and an improved track forecast. The track improvement at later times in the forecast with GH is primarily attributed to the change of the upper-level steering.

In the other case, Hurricane Gaston (2016) at 1200 UTC 27 August was in a favorable environment when a mid-to-upper-level low in the southeast began to move away from the storm. A butterfly-shaped sampling pattern was utilized in this mission. The GH dropsondes during the 6-h data window mainly come from the flight legs located in the southeast near-environment region of the TC. The assimilation of the GH dropsondes resulted in a noticeable improvement in the initial-time position (Fig. 14d), likely due to better definition of the mid-to-upper-level low in the near environment. The TC in both forecasts was then steered in similar environments (Figs. 14d,e). By 24–48 h, a blocking ridge started to build to the northwest of the TC. The run with the GH dropsonde data had a delayed northeastward turn due to being farther west than the other forecast, embedding it near this blocking ridge; the forecast without the dropsonde data turned northeastward away from the

influence of the ridge around 24 h. The TC in the forecast with GH data was then positioned farther north, compared to the one in the forecast without GH. Thus, the TC in the forecast with GH traveled northeastward much faster in the midlatitude westerlies along with slightly amplified ridges in the southwest and southeast of the TC (Fig. 14f). The improved track forecast in this case is largely due to an improved initial-time position from the assimilation of GH dropsondes.

8. Summary and conclusions

Results from a detailed composite study to investigate GH dropsondes impacts on TC analyses and forecasts in a vortex-scale ensemble DA system (HEDAS) and the HWRF Model are presented. Innovation statistics of assimilated dropsondes are evaluated to ensure that data are properly assimilated. The probability distribution functions of dropsonde temperature, wind, and humidity innovations indicate less bias and smaller root-mean-square error after DA.

Since reconnaissance data from crewed aircraft missions have shown a large impact on accurately defining

TC structure and the resulting forecasts (e.g., Christophersen et al. 2017), GH dropsonde data denial experiments are performed in two sets: one baseline set with crewed aircraft data and one without. The majority of the composite analyses are evaluated using the samples without crewed aircraft data. In this scenario, experiments with GH dropsondes are generally associated with smaller position and intensity errors and a better wind–pressure relationship in the final mean analyses. Better TC structures are also inferred from smaller integrated kinetic energy (IKE) errors in the final mean analyses when GH dropsondes are assimilated. Overall, assimilation of GH dropsonde data also results in statistically significant improvements of $\sim 10\%$ in track forecasts through 72 h and MSLP forecasts through 108 h.

Among the more intense TC cases (≥ 50 kt), two subgroups are considered based on 24-h intensity change: steady-state (SS) and nonsteady state (non-SS). GH dropsondes demonstrate a larger impact on the non-SS cases than the SS cases. This impact is evident in terms of a large reduction of position, intensity, and IKE errors during the 4-h DA window, as well as superior relative skill for position, intensity, and structure in the final mean analyses. As a consequence, the resulting forecasts for non-SS cases are generally associated with higher skill for both track and intensity, compared to SS cases. It is noteworthy that the non-SS track forecast skill is higher than those of all cases.

The forecast impact of GH dropsondes in the presence of crewed aircraft data is also examined. Overall, statistically significant track-forecast improvements are seen with GH dropsondes regardless of whether crewed aircraft data were available, but intensity improvements from GH dropsondes are only seen (beyond 36 h) when crewed aircraft data were absent. When assimilated with crewed aircraft data, GH dropsondes have a mixed impact on intensity forecasts.

We note that some statistically significant degradations of the composite intensity forecast for the non-SS cases between 24 and 48 h and for cases with the presence of the crewed aircraft data at 12 and 84 h are likely due to small sample size. Hence, variability in the individual cases makes it difficult to draw broader conclusions and develop dynamic explanations for the impact of assimilating GH dropsondes. Regardless, GH dropsondes demonstrate great value for TC analyses and forecasts, particularly for TCs that undergo large intensity changes [≥ 20 kt (24 h) $^{-1}$]. Greater intensity improvements in the absence of data from crewed reconnaissance aircraft reveal the strengths of long-endurance UAS platforms to sample TCs that traditional crewed aircraft cannot reach due to range limitations. When traditional crewed aircraft data are present, the impact from GH dropsondes is smaller. This leads to the conclusion that in such

circumstances, it is critical to optimally coordinate sampling strategies for crewed aircraft and UAS missions to exploit each aircraft's strengths. We further note that although the current study is an extension of the case studies presented by Christophersen et al. (2017), sample size in the current study is not sufficient to represent the characteristics of the two cases discussed in Christophersen et al. (2017) (i.e., not sufficient cases with double eyewall wind structure). Hence, we are not able to draw any general conclusions specifically related to these case studies.

Finally, it is noteworthy to clarify that the GFS ensemble-based analyses used as the initial and boundary conditions in this study did not assimilate GH dropsondes operationally; thus, the impact of GH dropsondes on lateral forcing is unknown with the current dataset. As GH dropsondes became operationally included in the GFS in 2017, future experiments should be devised to examine the full impact of GH dropsondes on regional and global model analyses and forecasts together. We further caution that generalizations drawn from results presented in this study must be made with restraint due to possible dependence of outcomes on the specifics of the DA and modeling systems employed. Moreover, in addition to the GH dropsonde observations that were the focus of the current study, the GH has carried various remote sensing instruments during NASA's HS3 and NOAA's SHOUT field campaigns. These observing platforms include the High-Altitude Monolithic Microwave Integrated Circuit (MMIC) Sounding Radiometer (HAMSR; Brown et al. 2011), the Scanning High-Resolution Interferometer Sounder (S-HIS; Revercomb and Taylor 2017), High-Altitude Imaging Wind Rain Airborne Radar (HIWRAP; Heymsfield et al. 2013) conically scanning Doppler radar, and cloud physics lidar (CPL; McGill et al. 2002). Assimilation of data from these instruments could modify the relative impacts of GH dropsondes, as well as observations from crewed aircraft in ways that may not be directly extrapolated from current findings.

Acknowledgments. The authors acknowledge funding support from the NOAA UAS Program (Award NA14OAR4830172) and from the NOAA Hurricane Forecast Improvement Project (HFIP) that also provided the computing resources. The data used in this study were collected by the NASA HS3 field campaign (available at https://espo.nasa.gov/hs3/content/HS3_0) and SHOUT campaign (available at ftp://ftp.aoml.noaa.gov/pub/hrd/data/global_hawk/). This research was carried out under the auspices of the Cooperative Institute for Marine and Atmospheric Studies, a joint institute of the University of Miami and NOAA (Cooperative Agreement NA67RJ0149). The authors thank constructive suggestions from Drs. Jonathan Poterjoy and Jason

Sippel and three anonymous reviewers that helped improve the quality of the manuscript.

REFERENCES

- Aberson, S. D., and J. L. Franklin, 1999: Impact on hurricane track and intensity forecasts of GPS dropwindsonde observations from the first-season flights of the NOAA Gulfstream-IV jet aircraft. *Bull. Amer. Meteor. Soc.*, **80**, 421–428, [https://doi.org/10.1175/1520-0477\(1999\)080<0421:IOHTAI>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<0421:IOHTAI>2.0.CO;2).
- , and C. Sampson, 2003: On the predictability of tropical cyclone tracks in the Northwest Pacific basin. *Mon. Wea. Rev.*, **131**, 1491–1497, [https://doi.org/10.1175/1520-0493\(2003\)131<1491:OTPOTC>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<1491:OTPOTC>2.0.CO;2).
- , A. Aksoy, T. Vukicevic, K. J. Sellwood, and X. Zhang, 2015: Assimilation of high-resolution tropical cyclone observations with an ensemble Kalman filter using HEDAS: Evaluation of the 2008–11 HWRf forecasts. *Mon. Wea. Rev.*, **143**, 511–523, <https://doi.org/10.1175/MWR-D-14-00138.1>.
- , K. J. Sellwood, and P. A. Leighton, 2017: Calculating dropwindsonde location and time from TEMP-DROP messages for accurate assimilation and analysis. *J. Atmos. Oceanic Technol.*, **34**, 1673–1678, <https://doi.org/10.1175/JTECH-D-17-0023.1>.
- Aksoy, A., 2013: Storm-relative observations in tropical cyclone data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **141**, 506–522, <https://doi.org/10.1175/MWR-D-12-00094.1>.
- , D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Wea. Rev.*, **138**, 1273–1292, <https://doi.org/10.1175/2009MWR3086.1>.
- , S. Lorsolo, T. Vukicevic, K. J. Sellwood, S. D. Aberson, and F. Zhang, 2012: The HWRf Hurricane Ensemble Data Assimilation System (HEDAS) for high-resolution data: The impact of airborne Doppler radar observations in an OSSE. *Mon. Wea. Rev.*, **140**, 1843–1862, <https://doi.org/10.1175/MWR-D-11-00212.1>.
- Braun, S. A., and Coauthors, 2013: NASA's Genesis and Rapid Intensification Processes (GRIP) field experiment. *Bull. Amer. Meteor. Soc.*, **94**, 345–363, <https://doi.org/10.1175/BAMS-D-11-00232.1>.
- , P. A. Newman, and G. M. Heymsfield, 2016: NASA's Hurricane and Severe Storm Sentinel (HS3) investigation. *Bull. Amer. Meteor. Soc.*, **97**, 2085–2102, <https://doi.org/10.1175/BAMS-D-15-00186.1>.
- Brown, S. T., B. Lambrigtsen, R. F. Denning, T. Gaier, P. Kangaslahti, B. H. Lim, J. M. Tanabe, and A. B. Tanner, 2011: The High-Altitude MMIC Sounding Radiometer for the Global Hawk unmanned aerial vehicle: Instrument description and performance. *IEEE Trans. Geosci. Remote Sens.*, **49**, 3291–3301, <https://doi.org/10.1109/TGRS.2011.2125973>.
- Chou, K.-H., C.-C. Wu, P.-H. Lin, S. D. Aberson, M. Weissmann, F. Harnisch, and T. Nakazawa, 2011: The impact of dropwindsonde observations on typhoon track forecasts in DOTSTAR and T-PARC. *Mon. Wea. Rev.*, **139**, 1728–1743, <https://doi.org/10.1175/2010MWR3582.1>.
- Christophersen, H., A. Aksoy, J. Dunion, and K. J. Sellwood, 2017: The impact of NASA Global Hawk unmanned aircraft dropwindsonde observations on tropical cyclone track, intensity, and structure: Case studies. *Mon. Wea. Rev.*, **145**, 1817–1830, <https://doi.org/10.1175/MWR-D-16-0332.1>.
- Cione, J. J., E. A. Kalina, E. W. Uhlhorn, A. M. Farber, and B. Damiano, 2016: Coyote unmanned aircraft system observations in Hurricane Edouard. *Earth Space Sci.*, **3**, 370–380, <https://doi.org/10.1002/2016EA000187>.
- DeMaria, M., C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398, <https://doi.org/10.1175/BAMS-D-12-00240.1>.
- Doyle, J., and Coauthors, 2017: A view of tropical cyclones from above: The Tropical Cyclone Intensity experiment. *Bull. Amer. Meteor. Soc.*, **98**, 2113–2134, <https://doi.org/10.1175/BAMS-D-16-0055.1>.
- Evans, C., and Coauthors, 2017: The extratropical transition of tropical cyclones. Part I: Cyclone evolution and direct impacts. *Mon. Wea. Rev.*, **145**, 4317–4344, <https://doi.org/10.1175/MWR-D-17-0027.1>.
- Gamache, J. F., 2005: Real-time dissemination of hurricane wind fields determined from airborne Doppler radar. JHT Project Final Rep., 38 pp., http://www.nhc.noaa.gov/jht/2003-2005reports/DOPLRGamache_JHTfinalreport.pdf.
- Gopalakrishnan, S. G., F. Marks, J. A. Zhang, X. Zhang, J.-W. Bao, and V. Tallapragada, 2013: A study of the impacts of vertical diffusion on the structure and intensity of the tropical cyclones using the high-resolution HWRf system. *J. Atmos. Sci.*, **70**, 524–541, <https://doi.org/10.1175/JAS-D-11-0340.1>.
- Hamill, T. M., J. S. Whitaker, D. T. Kleist, M. Fiorino, and S. J. Benjamin, 2011: Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods. *Mon. Wea. Rev.*, **139**, 3243–3247, <https://doi.org/10.1175/MWR-D-11-00079.1>.
- Heymsfield, G. M., L. Tian, L. Li, M. McLinden, and J. I. Cervantes, 2013: Airborne radar observations of severe hailstorms: Implications for future spaceborne radar. *J. Appl. Meteor. Climatol.*, **52**, 1851–1867, <https://doi.org/10.1175/JAMC-D-12-0144.1>.
- Hock, T. F., and J. L. Franklin, 1999: The NCAR GPS dropwindsonde. *Bull. Amer. Meteor. Soc.*, **80**, 407–420, [https://doi.org/10.1175/1520-0477\(1999\)080<0407:TNGD>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<0407:TNGD>2.0.CO;2).
- Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC 22, 21 pp., <https://www.nhc.noaa.gov/pdf/NWS-NHC-1988-22.pdf>.
- Jones, S. C., and Coauthors, 2003: The extratropical transition of tropical cyclones: Forecast challenges, current understanding, and future directions. *Wea. Forecasting*, **18**, 1052–1092, [https://doi.org/10.1175/1520-0434\(2003\)018<1052:TETOTC>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<1052:TETOTC>2.0.CO;2).
- Kuo, Y.-H., T. K. Wee, S. Sokolovskiy, C. Rocken, W. Schreiner, D. Hunt, and R. A. Anthes, 2004: Inversion and error estimation of GPS radio occultation data. *J. Meteor. Soc. Japan*, **82**, 507–531, <https://doi.org/10.2151/jmsj.2004.507>.
- Majumdar, S. J., M. J. Brennan, and K. Howard, 2013: The impact of dropwindsonde and supplemental rawinsonde observations on track forecasts for Hurricane Irene (2011). *Wea. Forecasting*, **28**, 1385–1403, <https://doi.org/10.1175/WAF-D-13-00018.1>.
- McGill, M., D. Hlavka, W. Hart, V. S. Scott, J. Spinhirne, and B. Schmid, 2002: Cloud Physics Lidar: Instrument description and initial measurement results. *Appl. Opt.*, **41**, 3725–3734, <https://doi.org/10.1364/AO.41.003725>.
- Powell, M. D., and T. A. Reinhold, 2007: Tropical cyclone destructive potential by integrated kinetic energy. *Bull. Amer. Meteor. Soc.*, **88**, 513–526, <https://doi.org/10.1175/BAMS-88-4-513>.

- Rappaport, E. N., J.-G. Jiing, C. W. Landsea, S. T. Murillo, and J. L. Franklin, 2012: The Joint Hurricane Test Bed: Its first decade of tropical cyclone research-to-operations activities reviewed. *Bull. Amer. Meteor. Soc.*, **93**, 371–380, <https://doi.org/10.1175/BAMS-D-11-00037.1>.
- Revercomb, H. E., and J. K. Taylor, 2017: Hurricane and Severe Storm Sentinel (HS3) Scanning High-Resolution Interferometer Sounder (S-HIS). NASA Global Hydrology Center DAAC, Huntsville, AL, <https://doi.org/10.5067/HS3/SHIS/DATA201>.
- Rogers, R. F., and Coauthors, 2006: The Intensity Forecasting Experiment: A NOAA multiyear field program for improving tropical cyclone intensity forecasts. *Bull. Amer. Meteor. Soc.*, **87**, 1523–1538, <https://doi.org/10.1175/BAMS-87-11-1523>.
- , P. Reasor, and S. Lorsolo, 2013: Airborne Doppler observations of the inner-core structural differences between intensifying and steady-state tropical cyclones. *Mon. Wea. Rev.*, **141**, 2970–2991, <https://doi.org/10.1175/MWR-D-12-00357.1>.
- , J. A. Zhang, J. Zawislak, H. Jiang, G. R. Alvey III, E. J. Zipser, and S. N. Stevenson, 2016: Observations of the structure and evolution of Hurricane Edouard (2014) during intensity change. Part II: Kinematic structure and the distribution of deep convection. *Mon. Wea. Rev.*, **144**, 3355–3376, <https://doi.org/10.1175/MWR-D-16-0017.1>.
- Susskind, J., C. D. Barnett, and J. M. Blaisdell, 2003: Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Trans. Geosci. Remote Sens.*, **41**, 390–409, <https://doi.org/10.1109/TGRS.2002.808236>.
- Torn, R. D., 2014: The impact of targeted dropwindsonde observations on tropical cyclone intensity forecasts of four weak systems during PREDICT. *Mon. Wea. Rev.*, **142**, 2860–2878, <https://doi.org/10.1175/MWR-D-13-00284.1>.
- Uhlhorn, E. W., P. G. Black, J. L. Franklin, M. Goodberlet, J. Carswell, and A. S. Goldstein, 2007: Hurricane surface wind measurements from an operational stepped frequency microwave radiometer. *Mon. Wea. Rev.*, **135**, 3070–3085, <https://doi.org/10.1175/MWR3454.1>.
- Vömel, H., K. Young, and T. F. Hock, 2016: NCAR GPS dropsonde humidity dry bias. NCAR Tech. Note NCAR/TN-531+STR, 8 pp., <https://doi.org/10.5065/D6XS5SGX>.
- Watts, A. C., V. G. Ambrosia, and E. A. Hinkley, 2012: Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. *Remote Sens.*, **2012**, 1671–1692, <https://doi.org/10.3390/rs4061671>.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2).
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. International Geophysics Series, Vol. 59, Elsevier, 467 pp.
- Yuter, S. E., and R. A. Houze, 1995: Three-dimensional kinematic and microphysical evolution of Florida cumulonimbus. Part II: Frequency distributions of vertical velocity, reflectivity, and differential reflectivity. *Mon. Wea. Rev.*, **123**, 1941–1963, [https://doi.org/10.1175/1520-0493\(1995\)123<1941:TDKAME>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<1941:TDKAME>2.0.CO;2).
- Zawislak, J., H. Jiang, G. R. Alvey III, E. J. Zipser, R. F. Rogers, J. A. Zhang, and S. N. Stevenson, 2016: Observations of the structure and evolution of Hurricane Edouard (2014) during intensity change. Part I: Relationship between the thermodynamic structure and precipitation. *Mon. Wea. Rev.*, **144**, 3333–3354, <https://doi.org/10.1175/MWR-D-16-0018.1>.